Practical Reasoning with Attention Mechanisms

Massimo Cossentino
National Research Council of Italy
ICAR Institute, Palermo, Italy
massimo.cossentino@icar.cnr.it

Giovanni Pilato
National Research Council of Italy
ICAR Institute, Palermo, Italy
giovanni.pilato@icar.cnr.it

Guido Averna
National Research Council of Italy
ICAR Institute, Palermo, Italy
guido.averna@icar.cnr.it

Myrto Mylopoulos
Carleton University
Ottawa, Canada
myrto.mylopoulos@carleton.ca

John Mylopoulos
University of Toronto
Toronto, Canada
jm@cs.toronto.edu

***Abstract*—One of the most interesting questions in the field of artificial intelligent agents is whether it is feasible to design agent architectures that merge the classical emphasis on practical reasoning, with advanced psychological capabilities that are crucial for action guidance. This paper proposes an agent architecture, named CATALINA (Cognitive AgenT prActicaL reasonINg Architecture), that augments Bratman's classic BDI paradigm with attentional mechanisms, executive functions (Buehler 2022), and a global workspace (Baars 2005). The architecture is applied to a scenario illustrating how the collaboration of the Executive Functions and the Global Workspace enables the action-guiding behaviour of the agent.**

***Index Terms*—Cognitive architecture, Practical Reasoning, Attention, Consciousness, Agents, BDI.**

I. Introduction

Bratman's Belief-Desire-Intention (BDI) architecture [1] [2] is one of the most widely used approaches for building resource-bounded agents that have goals and reason about which ones to pursue with their next action. However, despite its widespread use, there have been no proposals that supplement the BDI architecture with attention mechanisms that shift the focus of the agent from one goal to another, based on incoming perceptual information and also account for consciousness and awareness. We are interested in extending the BDI architecture with two attention mechanisms, based on Baars' theory of attention/consciousness [3], [4] and Buehler's theory of the executive system for agent guidance [5], [6]. Our proposal constitutes a novel cognitive architecture that supports practical reasoning capabilities together with relevant attention mechanisms that make it practical and scalable.

The question of what mechanisms underlie consciousness is one of the most important in the field of cognitive science. [7]. Relatedly, one of the most pressing questions in Artificial Intelligence is if it is possible to design and implement architectures that support some of the functions that underlie consciousness [8] [9].

The field of robot consciousness research pursues a dual objective: developing robots capable of exhibiting forms of experiential and functional consciousness and using robots as tools to gain deeper insights into biological consciousness [10]. A step towards this direction is to investigate and design a software agent architecture provided with proper functions related to consciousness and awareness. There are several competing theories of consciousness, and so far, there has been no clear winner [11]. For the purposes of this paper, we remain neutral on this question.

Our long-term aim is to propose an agent architecture with specific features that implement key aspects of leading theories while improving Bratman's practical reasoner. In particular, in this first implementation step, we adopt insights from Baars' Global Workspace Theory (GWT) of attention/consciousness. To this, we include some additional functions inspired by Buehler's proposal that the executive system constitutes an agent's capacity to guide their actions. Therefore, in the current version of the proposed architecture, an agent is conscious of some information insofar as the agent attends to it, and that information enters into the global workspace, thus making it available to a range of psychological functions.

In detail, we are implementing mechanisms or functions that have been linked to (associated with) consciousness in one prominent theory, i.e., Baars' Global Workspace Theory. We are not claiming that we are implementing consciousness itself in the architecture (the architecture is not sentient). We would like to clarify our position regarding the selected theory (GWT): this is but one theory of consciousness among many existing proposals, and we do not wish to claim that it is the best theory of them all. We remain neutral on this. In future work, we plan to implement other consciousness-related mechanisms or functions, i.e., one that is responsible for higher-order representation.

Following these research directions, this paper presents CATALINA (Cognitive AgenT prActicaL reasonINg Architecture): an innovative agent architecture built on top of the BDI approach and embedding the *Global Workspace Theory* and the *Executive System Theory*. We believe that this integration can lead to more robust and adaptable BDI agents capable of complex reasoning and improved performance in uncertain situations while operating on limited computational resources.

We acknowledge the support of the PNRR project FAIR - Future AI Research (PE00000013), Spoke 9 - Green-aware AI, under the NRRP MUR program funded by the NextGenerationEU.

The proposed architecture can constitute the backbone of a robotic autonomous system being capable of making conscious and rational decisions while pursuing multiple goals simultaneously, even under time and resource constraints. This characteristic can be useful in many fields such as cognitive robotics, affective robotics, etc. The novelty of this paper goes beyond the simple integration of the three aforementioned frameworks. It consists in: *1)* Developing a conceptual framework that aligns key concepts (beliefs, desires, intentions, attention, etc.) to ensure coherence across modules; *2)* Designing algorithms to implement Global Workspace behaviours; *3)* Creating an attention modulation mechanism which exploits two thresholds (saliency and attention) to guide the agent's reasoning based on focus and salience of intentions; and *4)* Establishing an agent work cycle that supports the adopted consciousness model.

The rest of the paper is structured as follows. Section II provides the research baseline, section III illustrates the proposed architecture, and section IV reports an example of how CATALINA realizes the consciousness model and attention modulation mechanism. Conclusions and future work are presented in section V.

II. Research Baseline

The proposed architecture is based on three key elements: the Baar's Global Workspace Theory, the Buehler's Executive System functions, and the Bratman's Belief-Desire-Intention (BDI) model. In this section, we briefly summarize the theories and models that inspired the CATALINA architecture.

A. Belief-Desire-Intention model of Bratman

The BDI model of Bratman [2] is an architecture for agent practical reasoning. In this model, three concepts are the core of practical reasoning: *beliefs*, *desires*, and *intentions*. A belief represents the state of anything the agent knows about the world; its value is related to some real-world property and can change over time with that. A desire is a change in the state of something in the world that the agent wants to achieve. Not all desires may be pursued at the same time, so the agent deliberates about pursuing some of them by promoting them to intentions. Intentions are a set of actions (plans) that an agent voluntarily decides to enact to achieve the desired change in the state of the world.

Bratman's BDI model includes several components, including the *Means-End Reasoner*, the *Opportunity Analyzer*, the *Filtering Process*, and the *Deliberation Process*. The Means-End Reasoner uses the agent's beliefs and desires to retrieve an existing plan from its repository or to conceive a new one if needed. These plans constitute the options the agent has to fulfil its desires. Such options will be provided to the Filtering Process, which rejects all options that could clash with the intentions (and related plans) that are in execution but also permits a revision of the current decisions to meet the changes perceived in the environment. The Opportunity Analyzer elaborates on the agent's desires and is ready to catch opportunities arising from the state of the world to improve the current intentions or consider new ones. Finally, the Deliberation Process considers all the filtered and surviving options and deliberates one or more useful options for promotion to intentions. The selected intentions will then be enacted by executing the actions specified in their options.

B. The Baars' Global Workspace Theory

Baars's Global Workspace Theory [3], [4] plays a key role in the proposed architecture. It is a shared global memory where all the knowledge is stored. The *Global Workspace* (GW) forwards incoming information to the psychological functions that may be interested in it. In this sense, it allows an agent to focus attention on specific information. Using a metaphor, Baars refers to this as a spotlight that illuminates a portion of memory, leaving the rest dark and obscured. Only this bright portion of the memory is conscious. The knowledge contained in the conscious (enlightened) part of the GW is sent to the other mind's modules.

C. Buehler's Executive System

The third key element that we consider in our architecture is based on the Buehler's discussion of the executive system and its role in agentive guidance [5], [6]. According to this view, an agent's psychology includes an *executive system* that manages several subsystems or subfunctions assigned to different tasks. These subsystems are four: the *Executive Inhibition Function System*, the *Executive Switching Function System*, the *Executive Resource Allocation Function System* and the *Working Memory Maintenance Function System*.
The Executive Switching Function initializes the phase of desire development and allows to alternate attention between conscious (*endogenous*) stimuli and conscious (*exogenous*) stimuli. The Executive Inhibition Function System inhibits everything that can interfere with the agent's goal, playing, together with the Executive Switching Function System, a key role in attention control. The Executive Resource Allocation Function System performs every action useful for the satisfaction of the agent's intentions. Finally, the Executive Working Memory Maintenance Function System takes care of the transfer and maintenance of information in memory between the Long Term Memory and the Global Workspace.

III. The Proposed Architecture

The proposed architecture [1] (see Fig. 1) supports both consciousness awareness and the focus attention mechanism [5] implementing a complex interaction between the executive functions (mainly among the Switching, Inhibition and Reasoner Functions) that is mediated by the Global Workspace where the agent's attention is focused on the portion of knowledge related to the current intentions.

In the following, for clarity, we describe the specific meaning of the key concepts that constitute the pillars of our approach and the use we make of them. For these concepts,

[1]The current version of the architecture is available for download from the CATALINA GitHub repository: https://github.com/CATALINA-Architecture/CATALINA_Model

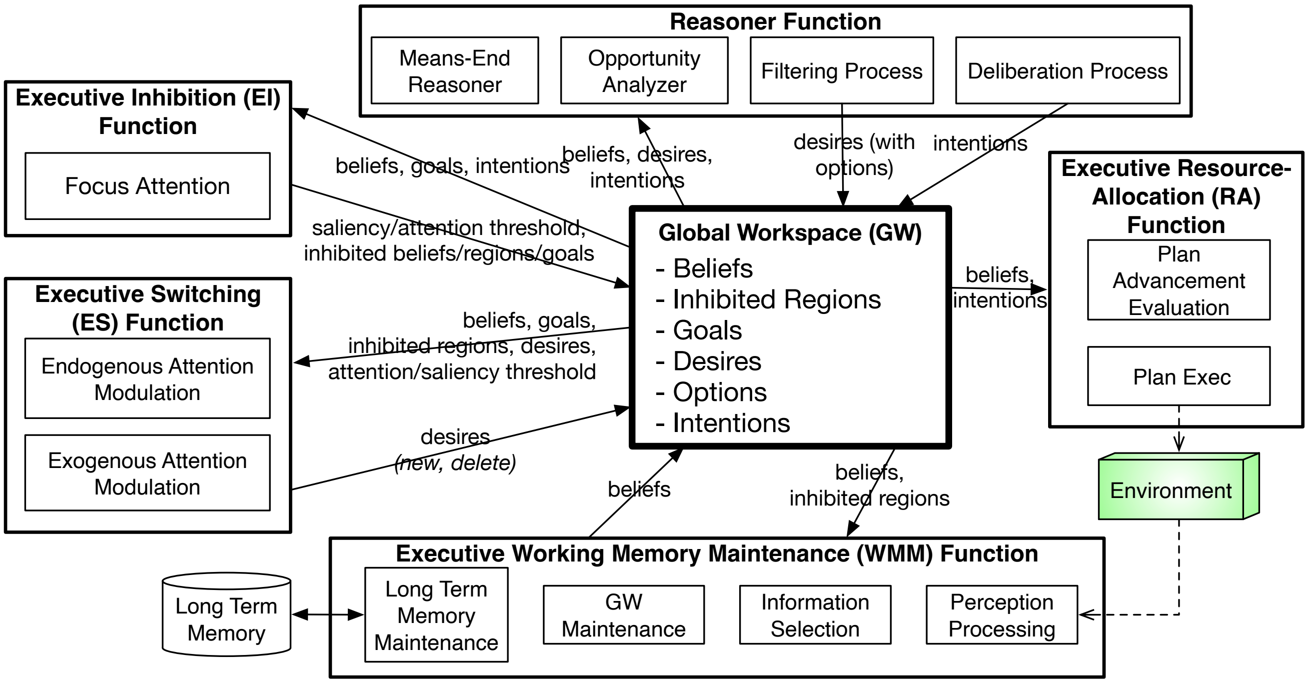

Fig. 1. The Proposed Cognitive AgenT prActicaL reasonINg Architecture (CATALINA)

we were inspired by Bratman's concepts [2], and one of our previous papers [12].

A **(functional) goal** is an intended state of the world; for instance, g_0:=*'Go To TownY'*. Usually, higher-level goals are decomposed into smaller ones, thus constituting a goal tree whose satisfaction is an agent's desire. When the entire goal tree cannot be satisfied, the agent may accept trade-offs that may result in the partial satisfaction of some goals (this functionality is not implemented in CATALINA yet). A goal also has a precondition defining the context in which it is to be pursued and a saliency expressing its relevance.

It is worth noting that when a goal is selected to be promoted as a desire, it becomes part of the set of wills of the agent. In fact, a **desire** is a goal that has passed the agent's saliency and attention thresholds filter. At the end of this process, the desire resides in the agent's mind as a will, but it is temporarily kept aside for future consideration.

The **Reasoner Function** evaluates the desires and generates one or more alternative plans (**options**) for each of them. Options may also come from the previous agent's experience. In fact, our architecture supports the repository we proposed in [13]. This trie-based repository supports the selection of plans according to quality attributes and their quick retrieval.

Finally, when the **Deliberation Process** selects the best option, the pair $(Desire, [Option])$ becomes a new intention that the Executive Resource-Allocation Function will execute. In other words, an **intention** is a desire the agent will pursue by enacting some optimal option that the agent puts into action. It is worth noting that an agent may also commit to an intention even if an option is not yet available for that intention. This allows the agent to focus its attention on the related part of the GW information and to avoid undertaking other intentions that could clash with this one.

There are other types of goals we explicitly consider in our architecture, more specifically: *epistemic* goals, *quality* goals, and the already cited *green* goals.

Epistemic goals are related to the need of the agent to update its knowledge, for instance: g_1:='*Is there ice on the road?*'. They may be triggered by a stimulus coming from some perception that motivates the agent to explore the environment. Epistemic goals are fulfilled by options (i.e. plans) exactly as functional goals are.

Quality goals represent the conventional concept of a quality property conditioning the actions of an agent; more specifically, a quality goal applies to a functional or epistemic goal and has a primary role in the selection of the best option for achieving that, when more than one is available. For instance, goal g_0 may be constrained by a quality goal: qg_0:= *'Drive safely'* implying, for instance, that the vehicle should drive well below the speed limits and further caution should be applied in case of rain. Quality goals participate in the trade-offs the agent will conceive to maximise the satisfaction of its goal tree. In fact, an agent may relax the quantitative constraint imposed by the quality goal, thus partially achieving its objective.

Finally, **green goals** represent the constraints the agent complies with in terms of environment respect policies during the pursuit of its goals. For instance gg_0 := '*Minimize* CO_2 *emissions*' may constrain goal g_0. These are a special kind of quality goal since they have some kind of normative legitimation that forbids the acceptance of any trade-off about their accomplishment.

According to our approach, the agent's knowledge is in the Global Workspace (as prescribed by Baars), in the form of **beliefs**. A belief includes a predicate and the credence the agent has about its truth state.

The behaviour of the agent's mind is implemented by the Executive Functions discussed by Buhelers [5], [14], namely: the Switching Function, the Inhibition Function, the Resource-Allocation Function, and the Working Memory Maintenance Function, and finally the Reasoner Function that supports some of the reasoning functionalities proposed by Bratman. Each Function is decomposed into several modules each one implementing specific portions of its behaviour.

In the next subsections, we will discuss these executive functions and their (sub-)modules.

A. The Global Workspace

The Global Workspace (GW) is a shared memory accessible to all executive functions within the proposed architecture. As described by Baar's GW theory, it is pivotal for implementing consciousness.

We illustrate here a working loop that starts from perception and arrives at action execution, showing the role of the GW.

The Executive Working Memory Maintenance (WMM) Function is responsible for transferring all pertinent information perceived by sensors into the GW. Raw data from sensors are processed, and the extracted information is stored within the GW as a belief update and made available to any function.

Specifically, the GW acts as a publish-subscribe dashboard; it processes incoming messages (such as belief updates) and generates outgoing events that notify of the update all the functions registered for that kind of information. Each Function can then access the GW and retrieve the updated belief when it needs it. In this way, the GW's operations do not interfere with the working process of each Function that operates autonomously, as proposed in Bueheler's theory.

Let us suppose a new belief representing the perception of some dangerous situation is posted in the GW (the Perception Processing and Information Selection modules cooperate to define the saliency of each perception). The Switching Function receives the notification of the availability of a new salient belief and retrieves that from the GW. The Switching Function compares its saliency with the current attention threshold: if the belief's saliency is greater than the saliency or attention threshold of the agent, then the Switching Function evaluates whether it is appropriate to generate a new (exogenous) epistemic goal that monitors the situation. The new goal is successively evaluated for promotion to desire so that the cause of the belief may be investigated. The Switching Function also revises the current set of desires, for instance, deleting one of them if some condition prevents its pursuit (for instance, precondition no more valid, clashing with current environment condition or other desires).

The new desire is processed by the Reasoner Function, which looks for new options and stores them in the GW. In the meanwhile, it reasons on the opportunity offered by the updated beliefs, and if this is the case, it revises the current options (options related to currently pursued intentions) and intentions (for instance, promoting a desire to intention).

The GW notifies the Resource Allocation Function of the updated beliefs and intentions so that this Function, when ready, may execute the new intentions while considering the updated state of the world. The actions generated by this Function will alter the environment, generating new perceptions that the Working Memory Maintenance Function will process; thus, the loop restarts.

B. The Executive Inhibition Function

The Executive Inhibition Function has two main duties: it implements the attention modulation mechanism and generates inhibition regions that limit the environment to areas where the agent focuses its perception. The attention modulation mechanism (implemented by the Focus Attention module) revises the saliency and attention thresholds. The inhibition regions are defined to focus the agent's attention on a specific portion of the environment; finally, this Function also defines the inhibition of beliefs that are not concerned with the current intentions, and therefore, their processing would slow the agent's mind.

C. The Executive Switching Function

In our architecture, we consider two different directions for attention propagation (and related modulation mechanisms): *(i)* Endogenous Attention Modulation is top-down attention related to a focusing effort guided by the agent's goals, *(ii)* Exogenous Attention Modulation is bottom-up attention driven by perceptual inputs that may overcome the inhibition barriers raised by the current attention threshold and require the agent to focus on a new stimulus that has higher saliency [15].

The Endogenous Attention Modulation process generates new agent desires. The agent has some goals and would like to pursue all of them, but in many cases, this is not practically possible, and therefore, only a few of them will be selected as intentions.

The Exogenous Attention Modulation process involves how attention is captured by external stimuli, like brightness or movement, which are processed through belief updates. This bottom-up attention may be automatic and involuntary.

When something meaningful happens, the agent directs its attention to processing the new stimulus, for instance, creating a new epistemic goal (and desire) to investigate the specific area of the environment that generated the perception.

When the Executive Switching Function receives an update of a belief from the GW, it evaluates whether the belief's saliency is greater than the current saliency/attention threshold. If it is, the agent considers whether the new perception requires revising the current set of desires (bottom-up attention modulation) thus adding new desires (with a higher saliency) or removing less important desires.

As already discussed, belief updates may trigger the bottom-up attention modulation mechanism that, if necessary, promotes some epistemic goal to a new desire aimed at exploring the origin of the perception that generated the new belief.

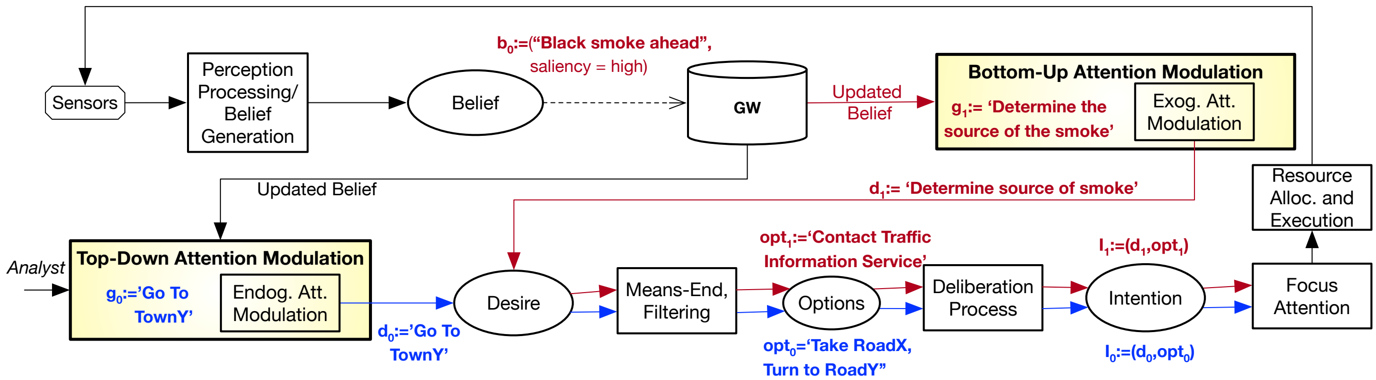

Fig. 2. The Attention mechanisms in CATALINA.

D. The Reasoner Function

The Reasoner Function is a higher-level function [14] inspired, in the CATALINA architecture, by the Bratman’s reasoner [2]. This function includes four modules: The first, the **Means-End Reasoner**, processes the new desires and searches the Plan Repository (in [13]) for plans that can satisfy them. If no suitable plan exists, a Planner sub-module is invoked to create a new plan and send it to the Filtering Process.

In parallel, the **Opportunity Analyser**, as suggested by Bratman [2], considers the current state of the world, the agent’s state, and looks for better opportunities to satisfy the agent’s current desires. For instance, it may find out, that in a specific situation, a good option arises for pursuing some desire that the agent is not currently acting upon. Suppose the agent is driving an autonomous vehicle from A to B but it also desires to pass from C (that is not in the best route connecting A to B). If some road blockage forces the vehicle to take another route, this may offer the opportunity to visit C without significant delay. This module is not implemented (yet) in the current version of CATALINA.

The **Filtering Process** evaluates the options’ quality attributes and publishes in the GW those that satisfy the quality and green desires of the agent.

Finally, the **Deliberation Process** decides which is the best option for satisfying each desire and if it may be adopted in the current agent’s state. This module decides if some agent’s desire may be actively pursued as an intention. The Deliberation Process also implements the trade-off capability, which consists of the algorithmic evaluation of degraded levels of qualities for the functional goals or even the relaxation of some parts of the goal formula (in terms of predicates or temporal constraints). It is worth noting that the current version of the CATALINA implementation does not yet support the trade-off capability, which is still under development.

E. The Executive Resource-Allocation (RA) Function

The Executive Resource Allocation (RA) Function executes the agent’s intention-embedded plan. This involves orchestrating the various actions that make up the plan and ensuring proper management, distribution, and equilibrium of resources. The current implementation performs the simple execution of the list of actions composing the plan, i.e. it invokes the corresponding agent’s behaviours. In future work, we plan to adopt some kind of workflow engine so that complex plans involving a parallel flow of actions may be supported.

F. The Executive Working Memory Maintenance (WMM) Function

The Executive Working Memory Maintenance (WMM) Function is responsible for managing, maintaining, and updating data within both the Long-Term Memory and the Global Workspace. It deals with sensor management addressing perceptions according to the prescriptions of the inhibited regions generated by the EI Function. This means that if a camera is looking at the road in front of an autonomous vehicle, this function removes from the processing area all the parts of the image that are inhibited (for instance, the sky that is not significant for driving the vehicle). Indeed, we have already discussed that bottom-up attention modulation allows for properly processing salient stimuli coming from the GW.

IV. An Example of Attention, Focusing and Consciousness

Buehler’s theory suggests that consciousness emerges from the interaction of executive functions like Switching and Inhibition, which enable flexible cognitive control and goal-directed behavior. On the other hand, the Baars’s Global Workspace Theory (GWT) can be seen as a complement of Buehler’s theory, since it describes consciousness as some sort of central hub that integrates and broadcasts information across cognitive processes. The integration of these theories can help to explain how the coordination of executive functions within the global workspace can support conscious awareness, complex decision-making, and deliberate action.

In this subsection, we will provide a quick example of how the proposed architecture realizes the consciousness model and the attention modulation mechanism we described above. The example refers to Fig. 2.

Let us suppose the agent, an autonomous vehicle (AV), starts in TownX with a goal g_0:=*‘Go To TownY by 6 pm’*.

This (functional) goal is constrained by the green goal gg_0 := '*Produce a low amount of* CO_2' and qg_0 := '*Drive safely*'. In Fig. 2, we can see that, initially, the Endogenous Attention Modulation module (of the Executive Switching Function) promotes the goal g_0 to desire and injects d_0 in the GW (the reader may follow the blue lines in the Figure). This desire is retrieved from the GW by the Means-End Reasoner of the Reasoner Function. Let us suppose that the Reasoner produces two different options. These are examined by the Filtering Process, which evaluates their respect for the quality goal qg_0 and green goal gg_0. Let us suppose the first option, opt_0, passes the filtering. The Deliberation Process considers the new desire, and since the agent is currently idle, it promotes d_0 plus the selected option opt_0 to intention I_0, and posts that to the GW. The Focus Attention module raises the saliency/attention thresholds and inhibits part of the GW knowledge thus representing the agent's willingness to focus on the selected intention. The Resource Allocation Function retrieves the new intention from the GW and starts executing that. The cycle closes by passing through the perception of the environment for monitoring the plan execution. Here, in Fig. 2, we suppose the agent perceives the presence of a dense black smoke far distant along the road. This is a relevant stimulus that is posted to the GW as a belief with a high saliency. The Exogenous Attention Modulation module of the Executive Switching Function retrieves the new belief from the GW and generates a new (epistemic) desire d_1 := '*Determine the source of the smoke*' with a very high saliency. This desire (red lines in the Fig. 2) is processed as we already described, and it generates a new intention (I_1) whose option consists of contacting a Traffic Information Service for evaluating the opportunity to take another route. This intention has a higher saliency than the previous one, and therefore, it is soon put into execution.

This portion of the agent's behaviour shows the collaboration among the executive functions and the central role of the GW in enabling that, coherently with the Buheler's theory of the Executive System and Baars' Global Workspace Theory.

As regards the focus attention behaviour that occurs when the agent deliberates a new intention, the Focus Attention module of the Inhibition Function decides:

- which areas of the environment should be considered less interesting for the current intentions (inhibition regions)
- which goals can not be any more promoted to desires (because they conflict with the current intentions)
- which beliefs (in the GW) should be considered relevant for the task at hand (enlightened by the spotlight) and which others should not be.

This involves the GW in the focusing behaviour exactly as proposed by Baars.

V. Conclusions and Future Work

We have illustrated an artificial intelligent agent architecture that we named CATALINA (Cognitive AgenT prActicaL reasonINg Architecture). The architecture supports *goal-oriented reasoning*, accepting *trade-offs* in pursuing multiple goals, providing the agent with *consciousness* features, which enable advanced reasoning. Furthermore, a *normative dimension* could condition the agent's decisions and trade-offs with a specific regard for green policies. The work takes inspiration from the Global Workspace Theory [4], thereby enabling an integrated implementation of key Executive Functions (specifically, Inhibitory Control, Resource Allocation, and Working Memory Maintenance) of the Executive System Theory [14].

Future work will consider using the proposed architecture in specific domains to determine its effectiveness and capabilities compared to other traditional approaches and experimentally validate the approach. Furthermore, we plan to enlarge the scope of this architecture to become a multi-tool that can support more than one approach to consciousness. In this direction, we plan to implement aspects of Rosenthal's higher-order theory of consciousness [16], [17], according to which an agent is in a conscious state in virtue of being aware of itself as being in that state by way of a higher-order representation, where this is typically accompanied by the ability to report the state that it is in. We have already done a preliminary study in that direction in [12], that we will deepen in the near future.

References

[1] M. Bratman, *Intention, Plans, and Practical Reason*. Harvard University Press, 1987.

[2] M. E. Bratman, D. J. Israel, and M. E. Pollack, "Plans and resource-bounded practical reasoning," *Computational intelligence*, vol. 4, no. 3, pp. 349–355, 1988.

[3] B. J. Baars, "The global workspace theory of consciousness: Predictions and results," *The Blackwell companion to consciousness*, pp. 227–242, 2017.

[4] B. J. Baars, "Global workspace theory of consciousness: toward a cognitive neuroscience of human experience," *Progress in brain research*, vol. 150, pp. 45–53, 2005.

[5] D. Buehler, "Agential capacities: a capacity to guide," *Philosophical Studies*, vol. 179, no. 1, pp. 21–47, 2022.

[6] D. Buehler, "Psychological agency-guidance of visual attention," Ph.D. dissertation, UCLA, 2014.

[7] M. Pokropski, *Mechanisms and consciousness: Integrating phenomenology with cognitive science*. Routledge, 2021.

[8] R. Arrabales, A. Ledezma, and A. Sanchis, "Criteria for consciousness in artificial intelligent agents," in *ALAMAS&ALAg Workshop at AAMAS 2008*, 2008, pp. 57–64.

[9] A. Chella, R. Manzotti *et al.*, "Artificial intelligence and consciousness," in *Association for the Advancement of Artificial Intelligence Fall Symposium*, 2007, pp. 1–8.

[10] A. Chella, A. Cangelosi, G. Metta, and S. Bringsjord, "Consciousness in humanoid robots," *Frontiers in Robotics and AI*, vol. 6, p. 17, 2019.

[11] A. Chella, "Robots and machine consciousness," in *Intelligent Robotics And Autonomous Agents Series - Cognitive Robotics*, A. Cangelosi and M. Asada, Eds. MIT press, 2022.

[12] M. Cossentino, L. Sabatucci, and J. Mylopoulos, "Consciousness as a trigger to adaptation," *Journal of Artificial Intelligence and Consciousness*, vol. 10, no. 01, pp. 27–47, 2023.

[13] M. Cossentino and G. Pilato, "Using a trie-based approach for storage and retrieval of goal-oriented plans in an s1/s2 cognitive architecture," *Cognitive Systems Research*, p. 101257, 2024.

[14] A. Diamond, "Executive functions," *Annual review of psychology*, vol. 64, no. 1, pp. 135–168, 2013.

[15] W. Wu, "We know what attention is!" *Trends in Cognitive Sciences*, vol. 28, no. 4, pp. 304–318, 2024.

[16] D. Rosenthal, "Higher-order awareness, misrepresentation and function," *Philosophical Transactions of the Royal Society B: Biological Sciences*, vol. 367, no. 1594, pp. 1424–1438, 2012.

[17] D. Rosenthal, "Consciousness and confidence," *Neuropsychologia*, vol. 128, pp. 255–265, 2019.